

Measuring Faculty EAP Self-Efficacy for Super Global Purposes: Using Rasch to Establish a Hierarchy of Faculty Developmental Needs

Aaron C. SPONSELLER

Institute for the Promotion of Global Education
Hiroshima University

Upon attaining the Japanese Ministry of Education, Culture, Science, and Technology's (MEXT) designation of the so-called "Super Global" university status, universities must necessarily shift their focus from merely setting aggressive goals to the actual implementation of policies and practices which will move them toward meeting those lofty goals. In many cases a significant portion of the benchmarks these universities are aiming to reach involve increasing the enrollment of foreign students, the number of foreign faculty, the courses taught in English, and the frequency with which their faculty and students present at international conferences and publish in international journals. The following research illustrates how one department at a major research university in Japan has taken the initial steps in this direction by developing and distributing a brief survey exploring faculty members' feelings of self-efficacy in English for academic purposes (EAP). By employing a Rasch-based analysis to validate the items on that survey, the department was then able to begin taking the initial steps towards meeting their Super Global goals by designing a series of workshops targeting the academic English skills identified by the faculty as most critically needed. The focus of the research presented here, however, is limited to a discussion of the challenges involved in the process of designing and validating an instrument capable of quickly measuring what faculty members believe they can and cannot do academically in English. Possible improvements for the next iteration of this instrument, which may be deployed across the entire university, are also included.

BACKGROUND

The Super Global Initiative

In their September 2014 press release announcing Type A and Type B universities, MEXT briefly summarized the "Top Global University Project" (referred to in this paper as "Super Global" hereafter) as follows:

The Top Global University Project is a funding project that aims to enhance the international compatibility and competitiveness of higher education in Japan. It provides prioritized support for the world-class and innovative universities that lead the internationalization of Japanese universities. Selected universities are expected to press forward with comprehensive internationalization and university reform. (MEXT, 2014, p.1)

Universities selected as either Type A or Type B submitted individual action plans. Hiroshima University, for example, named their plan the *Hiroshima University Global Campus Expansion and Innovation Initiative*.



FIGURE 1. Sample of SGU Benchmarks at Hiroshima University (Asahara & Sakakoshi, 2014, p.2)

Benchmarks to be reached by 2024 are at the heart of these plans. Many of the benchmarks set by universities appear to indicate the growing importance of English as the language of academia in Japan. (See Figure 1 for an example.)

While there is no specific mention of a particular language, it could be argued that use of English is absolutely necessary to even attempt meeting many of the benchmarks. Tripling the number of papers indexed by the Science Citation Index (SCI), tripling the amount of external funding, and teaching 50% of classes in a “foreign language” within ten years, as examples, clearly position English as a language vital to university success – inasmuch as success means hitting these benchmarks. University globalization efforts often boil down to embracing the ranking systems that quantify impact of academic activities. There may be pitfalls in these international ranking systems, but according to some experts such systems are an inevitability (Altbach, 2012) and the purpose of this paper is not to question the systems but to investigate how well prepared Japanese faculty members are to succeed under these system. So while the Super Global Initiative is not explicitly a language policy initiative, it can be argued that it represents a de facto language policy shift towards English.

English for Academic Purposes Self-Efficacy

EAP is a fairly broad term that applies to the English necessary to deal with academic language needs required at the university level (Brown, 2001). As Paltridge’s (2004) review on academic writing for second language learners illustrates, the number of potential skills nested under EAP writing alone is vast and necessarily mandates discussion of genre, communities of practice, corpus linguistics and vocabulary, and cross-cultural norms.

Perceived self-efficacy is commonly defined as “people’s beliefs about their capabilities to produce

designated levels of performance that exercise influence over events that affect their lives” (Bandura, 1994). Considering the university benchmarks shown in Figure 1 as a backdrop, it is clear that faculty members’ sense of self-efficacy in EAP skills is likely of critical importance to Japanese universities with Super Global status. Bandura’s (1977) initial research on self-efficacy has mushroomed considerably over the past four decades. Research on younger learners (Bandura, Barbaranelli, Caprara, & Pastorelli, 1996), academics (Pajares, 1996), and on aspects of second language learning (Eslami & Fatahi, 2008; Graham, 2011) have all shown it to be a critical psychological construct. Very little investigation into the self-efficacy of non-native English speaking academics who must use English in order to disseminate their research has been undertaken, however. The research presented here represents a small step forward in addressing this gap. For the purposes of the research conducted here, a comprehensive survey investigating self-efficacy of all possible EAP skills was not the point. The EAP skills investigated here were those most directly related to the Super Global benchmarks.

METHODS

Participants

Respondents ($n = 67$) were all members of the Faculty of Education at Hiroshima University. They completed the survey voluntarily and anonymously, thus they constitute a non-random convenience sample. Regarding age, 28.3% ($n = 19$) were 40 years of age or younger, 25.5% ($n = 17$) were 41 to 50 years old, 41.8% ($n = 28$) were 51 or over, and 4.5% ($n = 3$) declined to provide their age. Response by department is displayed in Table 1, below.

TABLE 1. Respondents by Department

Department	n	%
Elementary School Teacher Education	9	13.4
Special Needs School Teacher Education	6	9
Science Education	6	9
Mathematics Education	3	4.5
Technology and Information Education	1	1.5
Social Studies Education	4	6
Japanese Language and Culture Education	1	1.5
English Language and Culture Education	2	3
Teaching Japanese as a Second Language	3	4.5
Health and Sports Science Education	3	4.5
Human Life Science Education	6	9
Music Culture Education	2	3
Art Education	4	6
Educational Studies	8	11.9
Psychology	9	13.4
<i>Total</i>	67	100

Instrumentation

Self-efficacy items were written “*I can* [insert academic English task].” This is consistent with the approach advised by Bandura (2006). Each of the 12 items (as well as their Japanese equivalents) is listed in Figure 2 below in the order of their hypothesized hierarchy. Items were written in English and translated into

Japanese by a faculty support officer experienced in translation. Three highly proficient Japanese-English bilingual faculty members subsequently checked the items for clarity. Their suggested improvements, while minor, were incorporated into the final version of the survey items.

A six-point scale was used for the self-efficacy items. Response options were *strongly disagree* (全くそう思わない), *disagree* (そう思わない), *somewhat disagree* (あまりそう思わない), *somewhat agree* (ややそう思う), *agree* (そう思う), and *strongly agree* (強くそう思う). The decision to use scales without a neutral response option is in keeping with best practices in survey research in the field, no neutral response option was presented (Bond & Fox, 2013; Fulcher & Davidson, 2007; Nemoto & Beglar, 2014).

A paper version of the survey was provided to faculty members via their mailboxes in the faculty support office. A letter of appeal and thanks, in both Japanese and English, informed respondents about the purpose of the survey. A survey collection box was placed in the faculty support office. An electronic version of the survey was distributed via a link to a major online survey provider, Qualtrics (www.qualtrics.com).

Instrument Validation

The notion of validity has evolved over time. Fulcher and Davidson (2007) citing Cronbach and Meehl's (1955) early work, describe three "types" of validity: Criterion-oriented, content, and construct. Criterion-oriented validity refers to relationships between the instrument and the criterion to which test administrators are attempting to make inferences. Content validity essentially means the test is representative of the domain about which conclusions are being drawn. Construct validity involves defining what are frequently abstract notions in terms which make them observable and quantifiable. Constructs must

- | |
|---|
| <ol style="list-style-type: none"> 1. I can write a publishable academic paper in English.
私は独力でそのまま出版可能な英語論文を書くことができる。 2. I can give an academic presentation at a conference in English.
私は国際学会において英語での口頭発表を行うことができる。 3. I can speak English in diverse academic contexts.
私は研究分野と異なる学術的内容に関して英語で会話できる。 4. I can teach my current course(s) in English.
現在自分が開講している科目を英語でも講義できる。 5. I can provide academic guidance to international students in English.
私は留学生に英語で学術的な指導をできる。 6. I can write a syllabus for my course(s) in English.
英語のシラバスを簡単に作成できる。 7. I can give a poster presentation at a conference in English.
国際学会における英語でのポスター発表ができる。 8. I can write a presentation proposal for an international conference in English.
国際学会のための英文によるプロポーサルを容易に作成できる。 9. I can write an abstract in English.
英語の論文要旨を容易に書くことができる。 10. I can speak about my specific research interests in English.
私は関心のある研究分野に関して、英語で会話できる。 11. I can introduce myself to someone in English.
私は英語で自己紹介をできる。 12. I can write basic emails in English.
私は英文Eメールを作成できる。 |
|---|

FIGURE 2. Items Listed in Hypothesized Hierarchy: Most Difficult to Easiest

necessarily be defined in a manner which allows them to have relationships with other constructs. For a concept to be called a construct it must first be operationalized in a way that it can be observed and measured, otherwise it remains an abstract notion. Messick (1989) argued content and criterion validity were really part of the overall construct, and therefore construct validity is essentially king. Under his unified validity framework, he defined validity as “an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores or other modes of assessment” (Messick, 1989; as cited in Fulcher & Davidson, 2007).

Traditional models of test scoring lean heavily on raw scores which often fail to accurately reflect differences between individual test performances on an interval scale (Bond and Fox, 2013). The basic assumptions underlying the Rasch approach to validation are simple. Each person (e.g. test taker) has a level of ability. Similarly, not all test items are of equal difficulty, so each item on the test has a level of difficulty. The difference between person ability and item difficulty can be expressed along one line. This allows for establishment of probability estimates for the observation of correct or incorrect responses on any and all items by any individual person. Rasch analysis is a statistical procedure created for the sole purpose of instrument validation and investigating the differences between person ability and item difficulty.

Critical to the research presented here is the fact that the Rasch model also works well with polytomous/ Likert-style responses that do not generate right/wrong responses. The response scale and response scale category use statistics are key. In principle, a greater number of participants must complete Likert-style surveys in order to increase the robustness of Rasch analysis. The minimum number of responses in each response category is technically just one, but for robust analysis of Likert data at least 10 responses in each scale category is preferable (Wolfe & Smith, 2007; Linacre, 2004; Nemoto & Beglar, 2014).

RESULTS

Figure 3, a Wright Map, provides a graphic representation of person and item spread. Most respondents, represented by “x” marks on the left side of the map, are clustered lower on the scale indicating lower self-reported self-efficacy among the faculty overall. This is further confirmed in Table 2 which indicates Rasch-Andrich thresholds were adequately but not ideally spread, ranging from -1.38 to 1.36. Response category use statistics for all response options indicate adequate categorical use to generate interpretable data without necessitating the collapsing of any response categories. It is clear that response categories indicating a level of disagreement with items were far more popular than those indicating agreement. This again illustrates that faculty overall felt they lacked self-efficacy on most of the EAP items included on this survey.

In terms of technical quality, the procedure recommended by McNamara (1996) of using a value of ± 2 standard deviations (*SD*) of the Rasch infit mean-square (Infit MNSQ) statistic was used to determine which items were exhibiting the poorest fit. These twelve items, approached as a single construct, generated an infit MNSQ of .42. Therefore items falling outside of the infit MNSQ criterion rating of .16 (overfit) to 1.84 (underfit) were considered problematic. As the item-fit statistics presented in Table 3 indicate, one item (*I can speak English in diverse academic contexts*) just underfit the model, with an item infit MNSQ of 1.85. It is thus possible this item may have had a distorting impact on the measurement overall.

The substantive aspect of construct validity concerns the accuracy of the hypothesized item hierarchy (See Figure 2). The hypothesized hierarchy was not accurate outside of the relatively lower-order EAP tasks

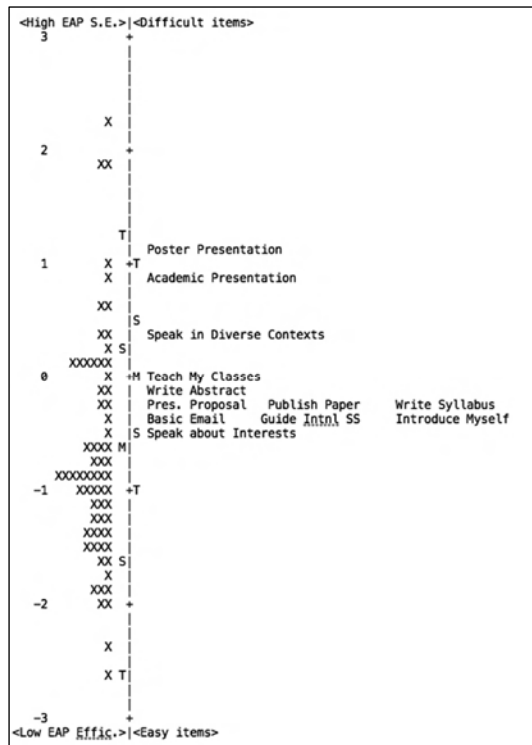


FIGURE 3. Item-Person Map (Wright Map) Displaying Actual Hierarchy of EAP Self-Efficacy

TABLE 2. Rating Scale Functioning for EAP Self-Efficacy

Response Category	Count (%)	Infit MNSQ	Outfit MNSQ	Andrich Thresholds	Category Measure
Strongly Disagree	182 (23)	1.15	1.12	(none)	(-2.81)
Disagree	184 (23)	.78	.86	-1.38	-1.38
Slightly Disagree	198 (25)	1.05	1.72	-.93	-.41
Slightly Agree	127 (16)	.98	1.14	.08	.43
Agree	65 (8)	.82	.77	.88	1.37
Strongly Agree	43 (5)	.54	.60	1.36	(2.78)

TABLE 3. Item Fit Statistics

Item	Measure	SE	Infit MNSQ	Infit ZSTD	Outfit MNSQ	Outfit ZSTD	Point-Measure Correlation
Poster Presentation	1.17	.16	1.70	3.0	2.69	5.2	.28
Academic Presentation	.92	.15	1.39	1.9	1.97	3.7	.35
Speak Diverse	.34	.13	1.85	4.0	1.71	3.3	.48
Teach Class	.01	.13	.58	-2.9	.59	-2.8	.87
Abstract	-.08	.12	.96	-.2	.93	-.4	.72
Presentation Proposal	-.21	.12	.86	-.9	.83	-1.0	.81
Syllabus	-.27	.12	.67	-2.2	.68	-2.1	.81
Publish a Paper	-.29	.12	.85	-.9	.85	-.9	.79
Introduce Myself	-.33	.12	.78	-1.4	.75	-1.6	.78
Guide Intn'l Students	-.36	.12	.63	-2.5	.63	-2.5	.81
Basic Email	-.41	.12	.56	-3.2	.58	-3.0	.82
Speak About Interests	-.49	.12	1.00	.0	1.04	.3	.70

of writing emails, introducing yourself, and speaking about your interests. This will be discussed in greater detail later. Related to the item hierarchy is the dimensionality of construct. The Rasch model accounted for 57.7% of the variance in the model (eigenvalue = 16.3). Of the unexplained variance, the first contrast had an eigenvalue of 3.1 and explained 10.8% of the variance in the model, indicating the probability that a second dimension was present in the model. The construct tentatively termed EAP Self-Efficacy for Super Global Purposes is likely not uni-dimensional.

Finally, regarding generalizability, the sample size was robust enough for reliable analysis of polytomous data with a six point response scale. The reliability of estimate statistic estimates the capacity of the test to accurately establish a hierarchy of items. The reliability of estimate will range from 0 to 1, and a higher value indicates a greater degree of confidence in the instrument. For this dataset item reliability (.92) and person reliability (.88) were relatively high.

DISCUSSION

The results reported above represent a first step in assessing the EAP self-efficacy of the Faculty of Education at Hiroshima University. The initial findings should be taken as preliminary at this time for several reasons which I will discuss here. First, and most problematic, is the suspicion that multiple constructs are present within the items. Second, the relatively small number of items on the instrument made deeper analysis of two constructs challenging. Third, a final hierarchy of items at this stage is not possible due to insufficient item/person spread. Fourth, it remains unclear if respondents were supplying answers based on experience or from their imagination. Finally, department-by-department analysis was not possible due to low response rates.

There is likely more than one construct at play in the data. Looking again at the Wright map (Figure 2), it seems clear that the first four items (*I can give a poster presentation in English, I can give an academic presentation in English, I can speak English in diverse academic contexts, and I can teach my classes in English*) are all face-threatening public speaking EAP skills. The next five skills are all writing skills (*I can write an abstract in English, I can write a presentation proposal in English, I can write a publishable paper in English, I can write a syllabus in English, and I can write a basic email in English*). Moreover, these five writing skills are very much clustered and exhibited very little separation whatsoever. The final three items may be conceptualized as low-stakes, non face-threatening EAP skills (*I can provide academic guidance for international students in English, I can write basic emails in English, and I can speak about my specific research interests in English*).

There was at least one item (*I can speak English in diverse academic contexts*) which performed problematically according to the infit MNSQ criterion. The word *diverse* would seem to be the primary culprit. While the item is not double-barreled syntactically, it may be so semantically. When respondents read the word “diverse” they may have imagined different scenes. Perhaps one person imagined a faculty meeting, another a cocktail party, and yet another a thesis defense. This item is problematic. It should be removed from future versions of the instrument and replaced with more specified items.

While all other items performed acceptably in a technical sense, there remains room for improvement. Two items, *I can write a publishable academic paper in English* and *I can teach my classes in English*, will be used as examples. Publishable academic papers can range from short reports written for in-house journals,

to opinion pieces, to cutting edge empirical research published in high impact factor journals. Teaching classes could mean several different things as well: undergraduate, graduate, lecture, seminar, omnibus, and so forth. Subsequent versions of this instrument should address these potential gaps by including more items which are more specific. This would help in at least two ways. First, the increased number of items would make it more possible to perform a robust analysis of multiple constructs if the initial data indicates the presence of multiple constructs (as was the case with this research). Second, the more specific the items the greater the chance is that the items will spread more adequately. Item spread for EAP writing skills was particularly problematic in this research.

It remains unclear which of the EAP skills the respondents had first-hand experience with and which they have never experienced. Bandura (1982) warns that self-efficacy judgments are related to action but that “faulty self-knowledge” and “misjudgment of task requirements” are two factors that can lead to discrepancies between self-efficacy and action (p.129). This is a critical point. Future versions of this instrument must incorporate some mechanism requesting respondents to indicate whether or not they have experience with the specific EAP skills the instrument addresses.

Finally, given the rather small response rate in many of the sub-faculties, department-by-department analysis was not possible. Future research in this area would do well to garner a more robust response from each faculty, as within-faculty analyses could potentially reveal different areas of EAP strength and weakness. This would, in turn, allow for the tailoring of faculty development workshops to best meet the needs of specific faculties.

CONCLUSION

In the most basic sense, “validity” in testing and assessment means the tests, assessments, protocols, or instruments used to make decisions are actually measuring what they claim to measure. Instrument and item validation is a process requiring thought on behalf of test creators or those who use test results in their decision-making. Statistical standards help guide instrument validation, but statistics are tools which *assist* the process rather than *guide* it. A conscientious test validator knows the construct and the theory underlying the construct, has defined and operationalized the construct with care, and hopes to confirm the construct is adequately measured through item measures, fit statistics, eigenvalues, residual loadings, and person/item reliability estimates.

The research presented here represents an initial attempt to measure the EAP self-efficacy of a faculty at a major Japanese research university with the Super Global University (SGU) designation. “Self-efficacy is not the only influence on behavior; it is not necessarily the most important... High self-efficacy will not produce competent performances when requisite skills are lacking” (Schunk, 1991, p.209). When skills are adequate and individuals expect positive outcomes that they also personally value, self-efficacy is thought to be a major influence on human behavior (Bandura, 1989). This is critical for major Japanese research universities seeking to globalize through the Super Global Initiative. The survey items were created using several of the benchmarks set by the university in its original application to become an SGU institution. Survey responses were subjected to a Rasch analytic approach, with the result being an initial hierarchy of EAP tasks/items in which the faculty, at-large report feeling more or less self-efficacious. Several critical issues remain unresolved, foremost being the probable presence of multidimensionality and the need to

include more items with greater specificity in the instrument. Doing so may allow for establishment of a more reliable item hierarchy with better item spread. Such a hierarchy would be invaluable for the purposes of organizing appropriate faculty development opportunities that focus upon the EAP skills requisite for meeting the university's benchmarks.

REFERENCES

- Altbach, P. G. (2012). The globalization of college and university rankings. *Change: The Magazine of Higher Learning*, 44(1), 26-31. <http://doi.org/10.1080/00091383.2012.636001>
- Asahara, T., & Sakakoshi, M. (2014). "Top Global University Project" Top Type: Hiroshima University's Global Campus Expansion and Innovation Initiative [Press Release]. Retrieved July 15, 2015 from http://s-en.hiroshima-u.jp/upload_files/download_files/Hiroshima-SGU.pdf
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, 84(2), 191.
- Bandura, A. (1989). Social cognitive theory. In R. Vasta (Ed.), *Annals of child development* (Vol. 6, pp. 1-60). Greenwich, CT: JAI.
- Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71-81). New York: Academic Press. (Reprinted in H. Friedman [Ed.], *Encyclopedia of mental health*. San Diego: Academic Press, 1998).
- Bandura, A. (2006). Guide for constructing self-efficacy scales. *Self-Efficacy Beliefs of Adolescents*, 5, 307-337.
- Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Multifaceted impact of self-efficacy beliefs on academic functioning. *Child development*, 1206-1222.
- Bond, T. G., & Fox, C. M. (2013). *Applying the Rasch model: Fundamental measurement in the human sciences*. Psychology Press.
- Brown, H. D. (2001). *Teaching by principles: An integrated approach to language pedagogy*. White Plains, NY: Pearson Education.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct Validity in Psychological Tests. *Psychological Bulletin*, 52(4), 281.
- Eslami, Z. R., & Fatahi, A. (2008). Teachers' Sense of Self-Efficacy, English Proficiency, and Instructional Strategies: A Study of Nonnative EFL Teachers in Iran. *TESL-EJ*, 11(4), n4.
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment*. Routledge London and New York.
- Graham, S. (2011). Self-efficacy and academic listening. *Journal of English for Academic Purposes*, 10(2), 113-117.
- Linacre, J. M. (2004). Optimal rating scale category effectiveness. In Smith, R. M. (2004). *Introduction to Rasch measurement: Theory, models and applications*. Jam Press.
- McNamara, T. F. (1996). *Measuring second language performance*. Addison Wesley Longman.
- MEXT (2014). *Selection for the FY 2014 Top Global University Project*. Tokyo: Higher Education Policy Planning Division, Higher Education Bureau, MEXT. Retrieved December 5, 2015 at http://www.mext.go.jp/b_menu/houdou/26/09/_icsFiles/afieldfile/2014/10/07/1_352218_02.pdf
- Messick, S. (1989). Validity. In Linn, R. L. (Ed.), *Educational measurement*. 3rd edition (p.13-103). New

York: Macmillan.

Nemoto, T., & Beglar, D. (2014). Likert-Scale Questionnaires. Retrieved from http://jalt-publications.org/files/pdf-article/jalt2013_001.pdf

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of educational research*, 66(4), 543-578.

Paltridge, B. (2004). Academic writing. *Language teaching*, 37(02), 87-105.

Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational psychologist*, 26(3-4), 207-231.

Wolfe, E. W., & Smith Jr, E. V. (2006). Instrument development tools and activities for measure validation using Rasch models: part II-validation activities. *Journal of Applied Measurement*, 8(2), 204-234.

ABSTRACT

Measuring Faculty EAP Self-Efficacy for Super Global Purposes: Using Rasch to Establish a Hierarchy of Faculty Developmental Needs

Aaron C. SPONSELLER

Institute for the Promotion of Global Education

Hiroshima University

Upon attaining the Japanese Ministry of Education, Culture, Science, and Technology's (MEXT) designation of the so-called "Super Global" university status, universities must necessarily shift their focus from merely setting aggressive goals to the actual implementation of policies and practices which will move them toward meeting those lofty goals. In many cases a significant portion of the benchmarks these universities are aiming to reach involve increasing the enrollment of foreign students, the number of foreign faculty, the courses taught in English, and the frequency with which their faculty and students present at international conferences and publish in international journals. The following research illustrates how one department at a major research university in Japan has taken the initial steps in this direction by developing and distributing a brief survey exploring faculty members' feelings of self-efficacy in English for academic purposes (EAP). By employing a Rasch-based analysis to validate the items on that survey, the department was then able to begin taking the initial steps towards meeting their Super Global goals by designing a series of workshops targeting the academic English skills identified by the faculty as most critically needed. The focus of the research presented here, however, is limited to a discussion of the challenges involved in the process of designing and validating an instrument capable of quickly measuring what faculty members believe they can and cannot do academically in English. Possible improvements for the next iteration of this instrument, which may be deployed across the entire university, are also included.

要 約

「スーパー・グローバル大学創生支援事業」下における 学部構成員の学術英語に対する自己効力感の調査 ー ラッシュ (Rasch, G.) の方法論を用いた学部構成員のための段階的な課題設定 ー

アーロン・C・スポンセラー

広島大学大学院教育学研究科

グローバル教育推進室

文部科学省が創設した「スーパー・グローバル大学創生支援事業」のもと、各大学はこれまで行ってきた大学のグローバル化のための方策や取り組みを、より高い目標へと転換する必要に迫られている。これまで、各大学において実施されてきたグローバル化を目指すための主要な取り組みは、留学生や外国人教員の増員、英語による授業の開講、教員や大学院生による国際学会発表や海外の学会誌への投稿の奨励などであった。本研究では、日本のある主要大学の一学部において、大学のグローバル化を図る手段として実施された、学部構成員の学術英語に対する自己効力感の調査について述べる。調査では、ラッシュ (Rasch, G.) の方法論を用いて学部構成員の学術英語能力に関する喫緊の課題を特定し、学部は、その結果に基づいて一連のワークショップの企画に着手した。本研究で提示した調査結果は、何をもって学部構成員が学術英語を使える／使えないと判断したかという点を含め、調査の過程と方法の両面において限界性を持つ。また、大学全体を対象として同じ調査を展開する場合にも、改良する余地を残すと思われる。